



```

QY 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120
DB 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120
QY 121 MNINQSAHGDRHGFIHARMRLPRKVVVGHWDREVRKIAKWMRVACAIQDGRGTQIVR 180
DB 121 MNINQSAHGDRHGFIHARMRLPRKVVVGHWDREVRKIAKWMRVACAIQDGRGTQIVR 180
QY 181 FGDNNREAVASTEDDKVEAQIKLGWSINTWVGSLAEGVKAVPENEVEELLKEYKERYIMP 240
DB 181 FGDNNREAVASTEDDKVEAQIKLGWSINTWVGSLAEGVKAVPENEVEELLKEYKERYIMP 240
QY 241 EDEYSLKAIREQAKMEIALREFLKEKNAIAFTTTTFEDLHDLPLQPLGLAVQRLMEGYGFG 300
DB 241 EDEYSLKAIREQAKMEIALREFLKEKNAIAFTTTTFEDLHDLPLQPLGLAVQRLMEGYGFG 300
QY 301 AEGDWKAAGLVRAKVMGAGLPGGTSFMEDYTYHLTPGNELVLGAHMLEVCPTIAKEKPR 360
DB 301 AEGDWKAAGLVRAKVMGAGLPGGTSFMEDYTYHLTPGNELVLGAHMLEVCPTIAKEKPR 360
QY 361 IEVHPLSIGGKADPARLVFDQEGPAPVNASIVDMGNRFRVLVNRVLSVPIERKMPKLP 420
DB 361 IEVHPLSIGGKADPARLVFDQEGPAPVNASIVDMGNRFRVLVNRVLSVPIERKMPKLP 420
QY 421 RVLWKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAEALEIYVLVIDENLDLENF 480
DB 421 RVLWKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAEALEIYVLVIDENLDLENF 480
QY 481 KKELRWNELYWGLLKR 496
DB 481 KKELRWNELYWGLLKR 496

```

```

RESULT 2
US-10-600-689-4
; Sequence 4, Application US/10600689
; Publication No. US20040058419A1
; GENERAL INFORMATION:
; APPLICANT: PYUN, Yu Ryang
; APPLICANT: KIM, Byoung Chan
; APPLICANT: LEE, Han Seung
; APPLICANT: LEE, Dong Woo
; APPLICANT: LEE, Yoon Hee
; TITLE OF INVENTION: THERMOSTABLE L-ARABINOSE ISOMERASE AND
; TITLE OF INVENTION: PROCESS FOR PREPARING D-TAGATOSE
; FILE REFERENCE: LEE881.001C1
; CURRENT APPLICATION NUMBER: US/10/600,689
; CURRENT FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: PCT/KR01/02243
; PRIOR FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: KR2000/80711
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: KR2000/80608
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Thermotoga neapolitana 5068
US-10-600-689-4

```

```

Query Match 100.0%; Score 2612; DB 15; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.7e-220; Indels 0; Gaps 0;
Matches 496; Conservative 0; Mismatches 0;

QY 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQASRIVEALNNDPIFPSPKIVLPVLSKSAEI 60
DB 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQASRIVEALNNDPIFPSPKIVLPVLSKSAEI 60
QY 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120

```

```

DB 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120
QY 121 MNINQSAHGDRHGFIHARMRLPRKVVVGHWDREVRKIAKWMRVACAIQDGRGTQIVR 180
DB 121 MNINQSAHGDRHGFIHARMRLPRKVVVGHWDREVRKIAKWMRVACAIQDGRGTQIVR 180
QY 181 FGDNNREAVASTEDDKVEAQIKLGWSINTWVGSLAEGVKAVPENEVEELLKEYKERYIMP 240
DB 181 FGDNNREAVASTEDDKVEAQIKLGWSINTWVGSLAEGVKAVPENEVEELLKEYKERYIMP 240
QY 241 EDEYSLKAIREQAKMEIALREFLKEKNAIAFTTTTFEDLHDLPLQPLGLAVQRLMEGYGFG 300
DB 241 EDEYSLKAIREQAKMEIALREFLKEKNAIAFTTTTFEDLHDLPLQPLGLAVQRLMEGYGFG 300
QY 301 AEGDWKAAGLVRAKVMGAGLPGGTSFMEDYTYHLTPGNELVLGAHMLEVCPTIAKEKPR 360
DB 301 AEGDWKAAGLVRAKVMGAGLPGGTSFMEDYTYHLTPGNELVLGAHMLEVCPTIAKEKPR 360
QY 361 IEVHPLSIGGKADPARLVFDQEGPAPVNASIVDMGNRFRVLVNRVLSVPIERKMPKLP 420
DB 361 IEVHPLSIGGKADPARLVFDQEGPAPVNASIVDMGNRFRVLVNRVLSVPIERKMPKLP 420
QY 421 RVLWKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAEALEIYVLVIDENLDLENF 480
DB 421 RVLWKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAEALEIYVLVIDENLDLENF 480
QY 481 KKELRWNELYWGLLKR 496
DB 481 KKELRWNELYWGLLKR 496

```

```

RESULT 3
US-10-193-896-11
; Sequence 11, Application US/10193896
; Publication No. US20030129710A1
; GENERAL INFORMATION:
; APPLICANT: Biotechnologisk Institut
; APPLICANT: Jorgensen, Flemming
; APPLICANT: Hansen, Ole C.
; APPLICANT: Stougaard, Peter
; APPLICANT: Berthelsen, Hans
; APPLICANT: Eriksnaer, Kristian
; APPLICANT: Bottcher, Karen
; APPLICANT: Christensen, Hans Jorgen Singel
; TITLE OF INVENTION: A novel thermostable isomerase and use
; TITLE OF INVENTION: hereof
; FILE REFERENCE: 30077US02
; CURRENT APPLICATION NUMBER: US/10/193,896
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/305,155
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/905,108
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 496
; TYPE: PRT
; ORGANISM: T. maritima
US-10-193-896-11

```

```

Query Match 96.0%; Score 2508; DB 14; Length 496;
Best Local Similarity 94.8%; Pred. No. 2.4e-211;
Matches 470; Conservative 18; Mismatches 8; Indels 0; Gaps 0;

QY 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQASRIVEALNNDPIFPSPKIVLPVLSKSAEI 60
DB 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQASRIVEALNNDPIFPSPKIVLPVLSKSAEI 60
QY 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120
DB 61 REIFEKANAEKPCAGVIVMMHTFSPSKMIRGLSINKKPLLHLHTQYNREIPWDTIDMDY 120

```

GénCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 22:42:57 ; Search time 467 Seconds  
(without alignments)  
1176.804 Million cell updates/sec

Title: US-10-600-689-4

Perfect score: 2612

Sequence: 1 MIDLKYBFWFLVGSQYLYG.....LENFKKELRWELNYGLLKR 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2612	100.0	496	27	US-10-193-896-12
2	2612	100.0	496	32	US-10-600-689-4
3	2508	96.0	496	27	US-10-193-896-11
4	2508	96.0	496	29	US-10-369-493-2914
5	2508	96.0	496	36	US-60-360-039-2914
6	1655	63.4	498	28	US-10-204-220-2
7	1653	63.3	497	27	US-10-193-896-6
8	1547	59.2	497	27	US-10-193-896-7
9	1547	59.2	497	29	US-10-369-493-17266
10	1547	59.2	497	36	US-60-360-039-17266
11	1456.5	55.8	597	30	US-10-446-203-12394
12	1443.5	55.3	500	27	US-10-193-896-4
13	1440.5	55.1	506	16	US-09-252-691-9419
14	1440.5	55.1	506	16	US-09-252-691C-9419
15	1440.5	55.1	506	30	US-10-417-886-9419
16	1436.5	55.0	500	29	US-10-369-493-698
17	1436.5	55.0	500	36	US-60-360-039-698
18	1434.5	54.9	500	27	US-10-193-896-3
19	1424.5	54.5	498	27	US-10-193-896-8
20	1424.5	54.5	498	29	US-10-369-493-23234
21	1424.5	54.5	498	36	US-60-360-039-23234
22	1315.5	50.4	501	27	US-10-193-896-13
23	1301	49.8	488	27	US-10-193-896-9
24	1270	48.6	488	27	US-10-193-896-10
25	1245	47.7	479	30	US-10-417-884-6017
26	1245	47.7	479	30	US-10-417-884A-6017
27	1212.5	46.4	498	27	US-10-193-896-5
28	389	14.9	465	24	US-09-905-108-2
29	389	14.9	465	27	US-10-193-896-2
30	364.5	14.0	407	27	US-10-193-896-14
31	310.5	11.9	108	23	US-09-865-590A-20004
32	137	5.2	471	15	US-09-103-611D-21
33	131	5.0	958	19	US-09-506-153-9
34	131	5.0	958	28	US-10-216-682-11
35	123.5	4.7	983	19	US-09-506-153-11
36	123.5	4.7	983	28	US-10-216-682-11
37	118.5	4.5	449	1	PCT-US03-13699-369
38	117	4.5	982	30	US-10-456-123-12
39	117	4.5	1481	22	US-09-791-537-87117
40	113.5	4.3	958	19	US-09-506-153-8
41	113.5	4.3	958	28	US-10-216-682-8
42	112	4.3	608	22	US-09-791-537-20913
43	112	4.3	608	28	US-10-282-122A-54351
44	111	4.2	949	30	US-10-456-123-11
45	109.5	4.2	925	30	US-10-425-115-231646

## ALIGNMENTS

RESULT 1  
US-10-193-896-12  
; Sequence 12, Application US/10193896  
; GENERAL INFORMATION:  
; APPLICANT: Biorekologisk Institut  
; APPLICANT: Jorgensen, Flemming  
; APPLICANT: Hansen, Ole C.  
; APPLICANT: Stougaard, Peter  
; APPLICANT: Berthelsen, Hans  
; APPLICANT: Eriknaer, Kristian  
; APPLICANT: Bottcher, Karen  
; APPLICANT: Christensen, Hans Jorgen Singel  
; TITLE OF INVENTION: A novel thermostable isomerase and use  
; FILE REFERENCE: 30077US02  
; CURRENT APPLICATION NUMBER: US/10/193,896  
; PRIOR FILING DATE: 2002-11-06  
; PRIOR APPLICATION NUMBER: 60/305,155  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/905,108

```

; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 496
; TYPE: PRT
; ORGANISM: T.neapol
US-10-193-896-12

Query Match      100.0%; Score 2612; DB 27; Length 496;
Best Local Similarity 100.0%; Pred. No. 6.3e-246;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQOASRIVEALNNDPIFSPKIVLKPVLKNSAEI 60
DB 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQOASRIVEALNNDPIFSPKIVLKPVLKNSAEI 60

QY 61 REIFEKANAEPKACAGIIVMMHTFSPSKMWIRGLSINKKPLLHLHTQYNNREIPWDITDMY 120
DB 61 REIFEKANAEPKACAGIIVMMHTFSPSKMWIRGLSINKKPLLHLHTQYNNREIPWDITDMY 120

QY 121 MNLNQSAGHGDREHGFTHARMRLPRKVVVGHWEDEVREKIAKWMRVACAIQDGRGTQIVR 180
DB 121 MNLNQSAGHGDREHGFTHARMRLPRKVVVGHWEDEVREKIAKWMRVACAIQDGRGTQIVR 180

QY 181 FGDNRREVASTEDDKVEAQIKLGWSINTWVGELAEAGVKA VPENEVEELLKEYKERYIMP 240
DB 181 FGDNRREVASTEDDKVEAQIKLGWSINTWVGELAEAGVKA VPENEVEELLKEYKERYIMP 240

QY 241 EDEYSLKAIKREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQPLGLAVQRLMEEGYGF 300
DB 241 EDEYSLKAIKREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQPLGLAVQRLMEEGYGF 300

QY 301 AEGDWKAAGLVRAKLVKMGAGLPGGTSFMEDYTHLTPGNELVLGAHMLVVCPTIAKEKPR 360
DB 301 AEGDWKAAGLVRAKLVKMGAGLPGGTSFMEDYTHLTPGNELVLGAHMLVVCPTIAKEKPR 360

QY 361 IEVHPLSIGGKADPARLVFDQGGPAPVNASIVDMGNRFLVNNRVLVSPIERKMPKLP 420
DB 361 IEVHPLSIGGKADPARLVFDQGGPAPVNASIVDMGNRFLVNNRVLVSPIERKMPKLP 420

QY 421 RVLMKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAELAEIEYLVIDENLDLENF 480
DB 421 RVLMKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAELAEIEYLVIDENLDLENF 480

QY 481 KKLERNELYWGLLKR 496
DB 481 KKLERNELYWGLLKR 496

RESULT 2
US-10-600-689-4
; Sequence 4, Application US/10600689
; GENERAL INFORMATION:
; APPLICANT: PYUN, Yu Ryang
; APPLICANT: KIM, Byoung Chan
; APPLICANT: LEE, Han Seung
; APPLICANT: LEE, Dong Woo
; APPLICANT: LEE, Yoon Hee
; TITLE OF INVENTION: THERMOSTABLE L-ARABINOSE ISOMERASE AND
; FILE REFERENCE: LEE81.001C1
; CURRENT APPLICATION NUMBER: US/10/600,689
; CURRENT FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: PCT/KR01/02243
; PRIOR FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: KR2000/80711
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: KR2000/80608
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```

; LENGTH: 496
; TYPE: PRT
; ORGANISM: Thermotoga neapolitana 5068
US-10-600-689-4

Query Match      100.0%; Score 2612; DB 32; Length 496;
Best Local Similarity 100.0%; Pred. No. 6.3e-246;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQOASRIVEALNNDPIFSPKIVLKPVLKNSAEI 60
DB 1 MIDLKQYEFWFLVGSQYLYGLETLLKVEQOASRIVEALNNDPIFSPKIVLKPVLKNSAEI 60

QY 61 REIFEKANAEPKACAGIIVMMHTFSPSKMWIRGLSINKKPLLHLHTQYNNREIPWDITDMY 120
DB 61 REIFEKANAEPKACAGIIVMMHTFSPSKMWIRGLSINKKPLLHLHTQYNNREIPWDITDMY 120

QY 121 MNLNQSAGHGDREHGFTHARMRLPRKVVVGHWEDEVREKIAKWMRVACAIQDGRGTQIVR 180
DB 121 MNLNQSAGHGDREHGFTHARMRLPRKVVVGHWEDEVREKIAKWMRVACAIQDGRGTQIVR 180

QY 181 FGDNRREVASTEDDKVEAQIKLGWSINTWVGELAEAGVKA VPENEVEELLKEYKERYIMP 240
DB 181 FGDNRREVASTEDDKVEAQIKLGWSINTWVGELAEAGVKA VPENEVEELLKEYKERYIMP 240

QY 241 EDEYSLKAIKREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQPLGLAVQRLMEEGYGF 300
DB 241 EDEYSLKAIKREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQPLGLAVQRLMEEGYGF 300

QY 301 AEGDWKAAGLVRAKLVKMGAGLPGGTSFMEDYTHLTPGNELVLGAHMLVVCPTIAKEKPR 360
DB 301 AEGDWKAAGLVRAKLVKMGAGLPGGTSFMEDYTHLTPGNELVLGAHMLVVCPTIAKEKPR 360

QY 361 IEVHPLSIGGKADPARLVFDQGGPAPVNASIVDMGNRFLVNNRVLVSPIERKMPKLP 420
DB 361 IEVHPLSIGGKADPARLVFDQGGPAPVNASIVDMGNRFLVNNRVLVSPIERKMPKLP 420

QY 421 RVLMKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAELAEIEYLVIDENLDLENF 480
DB 421 RVLMKPLPDFKRAATTAWILAGSSHHTAFSTAVDVEYLDWAELAEIEYLVIDENLDLENF 480

QY 481 KKLERNELYWGLLKR 496
DB 481 KKLERNELYWGLLKR 496

RESULT 3
US-10-193-896-11
; Sequence 11, Application US/10193896
; GENERAL INFORMATION:
; APPLICANT: Biotechnologisk Institut
; APPLICANT: Jorgensen, Flemming
; APPLICANT: Hansen, Ole C.
; APPLICANT: Stougaard, Peter
; APPLICANT: Berthelsen, Hans
; APPLICANT: Eriksner, Kristian
; APPLICANT: Botcher, Karen
; APPLICANT: Christensen, Hans Jorgen Singel
; TITLE OF INVENTION: A novel thermostable isomerase and use
; FILE REFERENCE: hereof
; FILE REFERENCE: 30077US02
; CURRENT APPLICATION NUMBER: US/10/193,896
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/305,155
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/905,108
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 496
; TYPE: PRT
; ORGANISM: T.maritima

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 22:37:36 ; Search time 42' Seconds  
(without alignments)  
1136.274 Million cell updates/sec

Title: US-10-600-689-4  
Perfect score: 2612  
Sequence: 1 MIDLKQYEFWFLVGSQYLXG.....LENFKKELRWNLVYGLLKR 496  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2508	96.0	496	2 B72398	L-arabinose isomer
2	1547	59.2	497	2 A83884	L-arabinose isomer
3	1445.5	55.3	500	2 AC0515	L-arabinose isomer
4	1443.5	55.3	500	1 ISEBAB	L-arabinose isomer
5	1436.5	55.0	500	1 ISEBAB	L-arabinose isomer
6	1435.5	55.0	500	2 B85488	L-arabinose isomer
7	1435.5	55.0	500	2 B90637	L-arabinose isomer
8	1424.5	54.5	498	2 C69587	L-arabinose isomer
9	1424.5	54.5	500	2 AE0274	L-arabinose isomer
10	1301	49.8	488	2 G97065	L-arabinose isomer
11	1270	48.6	488	2 C97065	L-arabinose isomer
12	137	5.2	471	2 B72314	hypothetical prote
13	127.5	4.9	475	2 H95997	probable L-arabino
14	117	4.5	1481	1 QZDOP3	pyrimidine synthet
15	115	4.4	467	2 AG1203	hypothetical prote
16	112	4.3	608	2 G81397	hsp90 family heat
17	106.5	4.1	1651	2 F88750	protein vit-6 [imp
18	106.5	4.1	1651	2 B43081	vitellogenin vit-6
19	106	4.1	328	2 F81005	probable proteinas
20	106	4.1	1613	2 G64488	reverse gyrase (in
21	105.5	4.0	320	2 F75063	hypothetical prote
22	105	4.0	2201	2 A32160	tenascin-C - human
23	105	4.0	4466	1 S17231	dynamin beta heavy
24	104.5	4.0	2242	2 A57541	pyrimidine synthet
25	104	4.0	328	2 H82025	probable proteinas
26	104	4.0	1246	2 JQ0406	hypothetical prote
27	103.5	4.0	815	2 E70021	3-hydroxyacyl-CoA
28	103	3.9	872	2 S76197	endopeptidase Clp
29	103	3.9	1152	2 AC1347	probable peptidogl

ALIGNMENTS

RESULT 1

B72398  
L-arabinose isomerase - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: B72398  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke-  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.  
C:M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se-  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: B72398  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-496 CARN>  
A:Cross-references: UNIPROT:Q9WYB3; GB:AE001709; GB:AE000512; NID:g4980763; PIDN:AAD353  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0276  
C:Superfamily: L-arabinose isomerase  
  
Query Match 96.0%; Score 2508; DB 2; Length 496;  
Best Local Similarity 94.8%; Pred. No. 4.3e-169;  
Matches 470; Conservative 18; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 MIDLKQYEFWFLVGSQYLXGLETLLKKVEQASRIVEALNNDPIFPSKIVLKPVLKNSAEI 60  
Db 1 MIDLKQYEFWFLVGSQYLXGLETLLKKVEQASRIVEALNNDPIFPSKIVLKPVLKNSAEI 60  
  
QY 61 REIFEKANAEKPCAGVIVMMHTFSPSKMWIRGLSINKKPLLLHLYTOYNEIPWDTTDMDY 120  
Db 61 TEIFEKANADPCAGVIVMMHTFSPSKMWIRGLSINKKPLLLHLYTOYNEIPWDTTDMDY 120  
  
QY 121 MNLNQSAGHDREHGFTHARMRLPRKVVVGHWDREVRKIAKMRVACAIQDGRGTQIVR 180  
Db 121 MNLNQSAGHDREHGFTHARMRLPRKVVVGHWDREVRKIAKMRVACAIQDGRGTQIVR 180  
  
QY 181 FGDNRREVASTDEDDKVEAQIKLGWSINTWGVGELARGVAVPENVEEILLKEYKEYIMP 240  
Db 181 FGDNRREVASTDEDDKVEAQIKLGWSINTWGVGELARGVAVPENVEEILLKEYKEYIMP 240  
  
QY 241 EDEYSIKATREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQGLAVORLMEEGYGF 300  
Db 241 EDEYSIKATREQAKMEIALREFLKEKNAIAFTTTFDLHDLPLQGLAVORLMEEGYGF 300  
  
QY 301 AEGDWKAAGLVRAIKVMGTSLPFGTSMEDYTYHLTPGNEVLGAHMLVCVPTIAKEKPR 360  
Db 301 AEGDWKAAGLVRAIKVMGTSLPFGTSMEDYTYHLTPGNEVLGAHMLVCVPTIAKEKPR 360  
  
QY 361 IEVHPLSIGGKADPARLPVFDGEGPAAVNASIVDMGNRFLVNRVLVSPIERKMPKLPTA 420  
Db 361 IEVHPLSIGGKADPARLPVFDGEGPAAVNASIVDMGNRFLVNRVLVSPIERKMPKLPTA 420

Db 361 IEVHPLSIGKADPARLVDFDQEGPANNASIVDMGNRFRVNVKVLSPRIKMPKLPPTA 420  
 Qy 421 RVLWKPLPDFKRATTANTILAGSHHTAFSTAVDVEYLIDWAEALIEIYLVIDENLDJENF 480  
 Db 421 RVLWKPLPDFKRATTANTILAGSHHTAFSTAVDVEYLIDWAEALIEIYLVIDENLDJENF 480  
 Qy 481 KKELRWNELYWGLLKR 496  
 Db 481 KKELRWNELYWGLLKR 496

RESULT 2  
 A:Accession: A83884  
 L-arabinose isomerase [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: A83884  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: A83884  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-497 <STO>  
 A:Cross-references: UNIPROT:Q9KBQ2; GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA0055  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: araA  
 C:Superfamily: L-arabinose isomerase

Query Match 59.2%; Score 1547; DB 2; Length 497;  
 Best Local Similarity 58.6%; Pred. No. 2.5e-101;  
 Matches 293; Conservative 72; Mismatches 127; Indels 8; Gaps 3;

Qy 1 MIDLKQYEFWFLVGSQVLYGLETLKKVQEQASRIVEALNNDPIFPKIVLKPVLKNSABI 60  
 Db 1 MLQTKPYTFWITGSOHLYGEDAIEQVROHSQTMVEKLNKIGELPYTIELKEVLTPDAI 60

Qy 61 REIFEFKANAEKPCAGVIVMHTFSPKMWIRGLSINKKPLHLHHTQVNRPIPDITDMDY 120  
 Db 61 RKMVIANSDDDCAGMITWHTFSPAKMWINGLQKPKLLHLHHTQVNRPIPDITDMDY 120

Qy 121 MNINQSAHGDREHGFTHARMRLPRKVVVGHVEDREVREKIAKMRVACAIQDGRGQIVR 180  
 Db 121 MNINQSAHGDREHGFTHARMRLPRKVVVGHVEDREVREKIAKMRVACAIQDGRGQIVR 180

Qy 181 FGDNMREVAIVEDGKVEAQIKLGSINTWVGELAGVKAQVENEVEBELLEKERY--- 237  
 Db 181 FGDNMREVAIVEDGKVEAQIKLGSINTWVGELAGVKAQVENEVEBELLEKERY--- 237

Qy 238 --IMPEDEYSLKAIQKAEALREFLEKNAIAFTTTFDLHDLPOLPGLAVORLME 295  
 Db 241 PSILEQDEVK-AAVLEQAKWELAKLEFLEGGYATFTNFEDLHGKQLPGLAVORLMAE 299

Qy 296 GYGFAGAGDWAAGLVRALKVNGAGLPGGTSFMEDYTYHLTPGNEVLGAHMLVCPTIA 355  
 Db 300 GYGFAGAGDWAAGLVRALKVNGAGLPGGTSFMEDYTYHLTPGNEVLGAHMLVCPTIA 357

Qy 356 KEKPRIEVHPLSIGKADPARLVDFDQEGPANNASIVDMGNRFRVNVKVLSPRIKMPKLPPTAR 415  
 Db 358 ANQPELOVHPLSIGKADPARLVDFDQEGPANNASIVDMGNRFRVNVKVLSPRIKMPKLPPTAR 417

Qy 416 KLPTARVLPDFKRAATTAWILAGSHHTAFSTAVDVEYLIDWAEALIEIYLVIDENLD 475  
 Db 418 KLPVAKVLWKCKPSLSEATEAMHAGGAHHTVFSFVTPQYQYDWTADLADIEVVPINDKT 477

Qy 476 DLENFKELRWNELYWGLLKR 495  
 Db 478 DVLQFOQQQLQWNEAFRRFLFK 497

RESULT 3

AC0515  
 L-arabinose isomerase [imported] - Salmonella enterica subsp. enterica serovar Typhi (set  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called salmonella typhi  
 C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AC0515  
 R:Farkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moulle, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AC0515  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-500 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD01259.1; PID:GL6501387; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY0119  
 C:Superfamily: L-arabinose isomerase

Query Match 55.3%; Score 1445.5; DB 2; Length 500;  
 Best Local Similarity 55.1%; Pred. No. 3.6e-94;  
 Matches 271; Conservative 82; Mismatches 134; Indels 5; Gaps 3;

Qy 7 YEFWFLVGSQVLYGLETLKKVQEQASRIVEALNNDPIFPKIVLKPVLKNSAEIREIFEK 66  
 Db 7 YEFWFLVGSQVLYGLETLKKVQEQASRIVEALNNDPIFPKIVLKPVLKNSAEIREIFEK 66

Qy 67 ANAEKPCAGVIVMHTFSPKMWIRGLSINKKPLHLHHTQVNRPIPDITDMDYNNLQNS 126  
 Db 67 ANYDRCAGLVVHLHTFSPAKMWINGLQKPKLLHLHHTQVNRPIPDITDMDYNNLQNS 126

Qy 127 AHGDRHGFTHARMRLPRKVVVGHVEDREVREKIAKMRVACAIQDGRGQIVRFGDNMR 186  
 Db 127 AHGDRHGFTHARMRLPRKVVVGHVEDREVREKIAKMRVACAIQDGRGQIVRFGDNMR 186

Qy 187 EVASTEDDKVEAQIKLGSINTWVGELAGVKAQVENEVEBELLEKERY-IMPEDY- 244  
 Db 187 EVASTEDDKVEAQIKLGSINTWVGELAGVKAQVENEVEBELLEKERY-IMPEDY- 244

Qy 245 --SLKAIQKAEALREFLEKNAIAFTTTFDLHDLPOLPGLAVORLMEEGVGGAE 302  
 Db 247 GDKQNVREARIEILGKWRFLQGGFRAFTTTFDLHDLPOLPGLAVORLMEEGVGGAE 306

Qy 303 GDWKAAGLVRALKVNGAGLPGGTSFMEDYTYHLTPGNEVLGAHMLVCPTIA-KEKPRI 361  
 Db 307 GDWKAAGLVRALKVNGAGLPGGTSFMEDYTYHLTPGNEVLGAHMLVCPTIA-KEKPRI 366

Qy 362 EVHPLSIGKADPARLVDFDQEGPANNASIVDMGNRFRVNVKVLSPRIKMPKLPPTAR 421  
 Db 367 DVQHLGIGGKEDPARLVDFDQEGPANNASIVDMGNRFRVNVKVLSPRIKMPKLPPTAR 426

Qy 422 VLWKPLPDFKRAATTAWILAGSHHTAFSTAVDVEYLIDWAEALIEIYLVIDENLDJENFK 481  
 Db 427 ALWKAQPDLPPTASEAWILAGSHHTAFSTAVDVEYLIDWAEALIEIYLVIDENLDJENFK 486

Qy 482 KKELRWNELYWGL 493  
 Db 487 DALRWNEVYGL 498

RESULT 4  
 ISEBAB  
 L-arabinose isomerase (EC 5.3.1.4) - Salmonella typhimurium  
 C:Species: Salmonella typhimurium  
 C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
 C:Accession: A24985  
 R:Lin, H.C.; Lei, S.P.; Wilcox, G.  
 Gene 34, 123-128, 1985  
 A:Title: The arabid operon of Salmonella typhimurium LT2. II. Nucleotide sequence of ara  
 A:Reference number: A24985; MUID:85232045; PMID:3891513